Make A Right Choice
-NAND Flash As Cache And Beyond

Simon Huang
Technical Product Manager

simon.huang@supertalent.com
Super Talent Technology

December, 2012

Release 1.01

www.supertalent.com
Legal Disclaimer

• INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH SUPER TALENT™ PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN SUPER TALENT’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, SUPER TALENT ASSUMES NO LIABILITY WHATSOEVER, AND SUPER TALENT DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF SUPER TALENT PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

• Super Talent may make changes to specifications and product descriptions at any time, without notice. Super Talent may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights. Designers must not rely on the absence or characteristics of any features or instructions marked “reserved” or “undefined.” Super Talent reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. Contact your local Super Talent sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Super Talent literature may be obtained by visiting Super Talent's website at http://www.supertalent.com. Super Talent™ is a trademark or registered trademark of Super Talent or its subsidiaries in the United States and other countries.

• Other names and brands may be claimed as the property of others
Where can we use NAND Flash?

- iPhone
- Embedded
- Netbook
- Android Phone
- Laptop/Ultrabook
- HPC
- Data Center
- Cloud Server
- Web Server
- Telecom
- IPTV
- iPad
- Embedded
- iPhone

Everywhere in Computing
## SSD Unit Shipment Forecast

**Worldwide SSD Unit Sales**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>46.9</td>
<td>54.4</td>
<td>98.1</td>
<td>139.6</td>
<td>193.2</td>
<td>258.5</td>
</tr>
<tr>
<td>Embedded</td>
<td>0.5</td>
<td>1.1</td>
<td>2</td>
<td>2.9</td>
<td>4.1</td>
<td>6.1</td>
</tr>
<tr>
<td>Client Dual</td>
<td>7.7</td>
<td>1.5</td>
<td>19.1</td>
<td>27.8</td>
<td>39.7</td>
<td>56.6</td>
</tr>
<tr>
<td>Client Single</td>
<td>15.9</td>
<td>26.1</td>
<td>48.7</td>
<td>78.2</td>
<td>116.3</td>
<td>160</td>
</tr>
<tr>
<td>Enterprise</td>
<td>22.8</td>
<td>25.7</td>
<td>28.3</td>
<td>30.7</td>
<td>33.1</td>
<td>35.8</td>
</tr>
</tbody>
</table>

Source: Objective Analysis Data, 2012
What is Cache?

- A cache is simply a copy of a small data segment residing in the main memory.
- Fast but small extra memory.
- Hold identical copies of main memory.
- Lower latency.
- Higher bandwidth.
- Usually several levels (1, 2 and 3).
Why cache is so important?

• Old days: CPUs clock frequency was the primary performance indicator.
• Microprocessor execution speeds are improving at a rate of 50%-80% per year while DRAM access times are improving at only 5%-10% per year.
• If the same microprocessor operating at the same frequency, system performance will then be a function of memory and I/O to satisfy the data requirements of the CPU.
There are three types of cache that are now being used:

- One on-chip with the processor, referred to as the "Level-1" cache (L1) or primary cache
- Another is on-die cache in the SRAM is the "Level 2" cache (L2) or secondary cache.
- L3 Cache

PCs and Servers, Workstations each use different cache architectures:

- PCs use an asynchronous cache
- Servers and workstations rely on synchronous cache
- Super workstations rely on pipelined caching architectures.
Typical Cache Configuration

- CPU
- L1 Register
- L1 Data Cache
- L1 Inst Cache
- L2 Cache
- L3 Cache
- Main Memory
How Cache is Used?

• Cache contains copies of some of Main Memory
  – those storage locations recently used
    • when Main Memory address A is referenced in CPU
    • cache checked for a copy of contents of A
  – if found, cache hit
    • copy used
    • no need to access Main Memory
  – if not found, cache miss
    • Main Memory accessed to get contents of A
    • copy of contents also loaded into cache
Why needs Cache?

- Due to increasing gap between CPU and main Memory, small SRAM memory called L1 cache inserted.

- L1 caches can be accessed almost as fast as the registers, typically in 1 or 2 clock cycle

- Due to even more increasing gap between CPU and main memory, Additional cache: L2 cache inserted between L1 cache and main memory: accessed in fewer clock cycles.
Why needs Cache (continue)?

- L2 cache attached to the memory bus or to its own cache bus.
- Some high performance systems also include additional L3 cache which sits between L2 and main memory. It has different arrangement but principle same.
- The cache is placed both physically closer and logically closer to the CPU than the main memory.
The HDD/NAND/DRAM Speed Gap

Bandwidth (MB/s)

Price per Gigabyte

Source: OBJECTIVE ANALYSIS

www.supertalent.com

NAND as Cache
CPU/Memory/NAND/HDD evolution

- **PicoSec**
- **NanoSec**
- **MicroSec**
- **MillSec**
- **Second**

- **Tape**
- **Hard Disk**
- **SATA/SAS SSD**
- **PCI-e SSD**
- **NAND Flash**
- **DRAM**
- **CPU**

- 1,000,000,000s
- 100,000,000 s
- 100,000s
- 100s per operation

www.supertalent.com
Keys to Consider using NAND

- Performance
- Capacity
- Bits per Cell
- Number of Write/Erase Cycles (Endurance)
- Data Retention
- Cost
- Cell Size/Lithography
Why NAND flash as cache so important?

• Increasing IOPS up to 20% to 30%
• Improving average response time up to 20%
• Less power up to 30% to 40%
• Lower storage cost up to 45% per TB
NAND flash Caching Architectures

Server

Network

Storage

Flash On Server
Closest to CPU
Lowest latency

Good for Cluster Servers

Flash on Storage Controller
Google Data Center

www.supertalent.com
NAND flash as Cache In Data Center

Server Level
- *PCI-e SSD on the Host
- *SAS/SATA SSD on the Host

Controller Level
- *Flash Cache

Disk Array Level
- *Flash Array Pool
- *Flash as Cache
# Cache Write Policy

<table>
<thead>
<tr>
<th>Policy</th>
<th>Write Back</th>
<th>Write Through</th>
<th>Write Around/Read only</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Store operation</strong></td>
<td>Write to SSD 1st then copy to HDD</td>
<td>Write to SSD and HDD at the same time</td>
<td>No write to SSD</td>
</tr>
<tr>
<td><strong>Data protection</strong></td>
<td>Data loss risk if write to SSD failure</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td>Middle</td>
<td>Low</td>
<td>High if Read Intensive</td>
</tr>
<tr>
<td><strong>Application</strong></td>
<td>Data Mining Searching</td>
<td>OLPT</td>
<td>Database/Web Searching</td>
</tr>
<tr>
<td><strong>STT Solutions</strong></td>
<td>TeraDrive/SuperNova SATA III SSD</td>
<td>TeraDrive/SuperNova SATA III SSD</td>
<td>TeraDrive/SuperNova SATA III SSD</td>
</tr>
</tbody>
</table>

- **Write**: SSD reads and writes are performed first, followed by HDD operations.
- **Read**: SSD reads are performed first, followed by HDD operations.
- **Write**: SSD writes are performed first, followed by HDD operations.
- **Read**: SSD reads are performed first, followed by HDD operations.
- **Read**: SSD reads are performed first, followed by HDD operations.
- **Write**: SSD writes are performed first, followed by HDD operations.
### Data Placement Strategy

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Primary Storage</th>
<th>Tiering Storage</th>
<th>Caching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity Usage</td>
<td>All</td>
<td>Frequently accessed Data</td>
<td>Copy of Frequently accessed Data</td>
</tr>
<tr>
<td>Data Protection</td>
<td>SSD failure cause data loss</td>
<td>SSD failure cause data loss</td>
<td>SSD failure impact operation a little</td>
</tr>
<tr>
<td>Write Policy</td>
<td>Read/Write Intensive</td>
<td>Read Intensive</td>
<td>Mixed Read/Write, Changing data access pattern</td>
</tr>
<tr>
<td>Application</td>
<td>Big data</td>
<td>Middle size data</td>
<td>A smaller chunk data</td>
</tr>
<tr>
<td>NAND Flash Type</td>
<td>SLC /eMLC/MLC</td>
<td>SLC/eMLC</td>
<td>SLC or eMLC</td>
</tr>
<tr>
<td>STT Solution</td>
<td>TeraDrive/SuperNova SATA III SSD</td>
<td>TeraDrive/SuperNova SATA III SSD/RAIDDRIVE II</td>
<td>TeraDrive/SuperNova SATA III SSD/RAIDDRIVE II</td>
</tr>
</tbody>
</table>

![Diagram showing data placement strategy]
# NAND Flash Type Comparison

<table>
<thead>
<tr>
<th>Type</th>
<th>P/E Cycle</th>
<th>Cost</th>
<th>Random Write Performance Comparing HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC</td>
<td>100k</td>
<td>High</td>
<td>5X</td>
</tr>
<tr>
<td>eSLC</td>
<td>50k</td>
<td>Middle High</td>
<td>3.75X</td>
</tr>
<tr>
<td>eMLC</td>
<td>30k</td>
<td>Middle</td>
<td>3X</td>
</tr>
<tr>
<td>MLC</td>
<td>10k</td>
<td>Low</td>
<td>2X</td>
</tr>
<tr>
<td>TLC</td>
<td>1K</td>
<td>Very Low</td>
<td>1X</td>
</tr>
</tbody>
</table>
What to expect your NAND flash device?

- **Data Retention**
- **ECC**
- **Controller & NANDs**
- **SSD and OSs**
- **SATA/SAS/PCI-e/PCI-e Express**
- **Wear Leveling**
- **Cache**
- **Overprovision**
- **Trim**
- **Boot Time**
- **Read/Write Speed**
- **IOPS**
- **Power Consumption**
- **Price**
- **Data Encryption: AES-128/256**
- **TCG Enterprise**
- **Reliability**
- **Security**
- **Compatibility**
- **Endurance**
Performance Tier for Enterprise Storage Systems

Tier 0
- Financial Transactions
- E-commerce Applications

PCI-E SSD
100K+ IOPS

Ultra high performance Enterprise Storage Systems

Tier 1
- Business Processing
- Data Analysis/Mining
- Cloud Computing
- Caching
- Data Centers

FC/SAS
HDD/SATA III
Extreme
IOPS SSD
50K+ IOPS

High performance Enterprise Storage Systems

Tier 2
- E-mail
- File and Print

SATA
HDD/SSD
25K+ IOPS

Low Cost HDD/SSD

Tier 3
- Data Backup
- Archive

TAPE/Offline

Lowest Cost Storage Media
NAND flash Solutions for Enterprise

- Server Based SSD has value for rapid boot
- PCIe has value for *caching* /storage memory
- Network Caching bring performance to legacy systems
- Storage Systems with integrated flash or flash only are compelling refreshes
Recap

• NAND Flash for Cache now is the critical part of the Server/Storage/Network
• Increase IOPS and lower IPOS/watt
• Cache Write policy and Data placement strategy impact IOPS and $ IPOS
• STT RAIDdrive, TeraNova and SuperNova are the right cache solution for Server/Storage/Network
Backup
<table>
<thead>
<tr>
<th>Architecture</th>
<th>System</th>
<th>Network</th>
<th>Technology</th>
<th>Component</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAS</td>
<td>Disk</td>
<td>Switch</td>
<td>FC</td>
<td>RAID Controller</td>
<td>OS</td>
</tr>
<tr>
<td>SAN</td>
<td>Tape</td>
<td>Directors</td>
<td>SAS</td>
<td>JBOD</td>
<td>Security</td>
</tr>
<tr>
<td></td>
<td>High End FC Array</td>
<td>Gateway/Bridge</td>
<td>SCSI</td>
<td>HBA</td>
<td>Deduplication</td>
</tr>
<tr>
<td>NAS</td>
<td>Mid End FC Array</td>
<td>Appliances</td>
<td>SATA</td>
<td>NIC/TOE</td>
<td>Virtualization</td>
</tr>
<tr>
<td>Hybrid</td>
<td>Unified Storage</td>
<td>iSCSI</td>
<td></td>
<td>NAS Head</td>
<td>Cloud Computing</td>
</tr>
<tr>
<td></td>
<td>Libraries</td>
<td>InfiniBand</td>
<td>iSCSI Head</td>
<td></td>
<td>Snapshot</td>
</tr>
<tr>
<td></td>
<td>Virtual Tape</td>
<td>GbE</td>
<td>CNA</td>
<td></td>
<td>Remote Duplication</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCIP</td>
<td></td>
<td></td>
<td>Thin Provision</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IFCP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCoE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Thank you

For more info:
Visit:
http://www.supertalent.com or
Email:
sales@supertalent.com